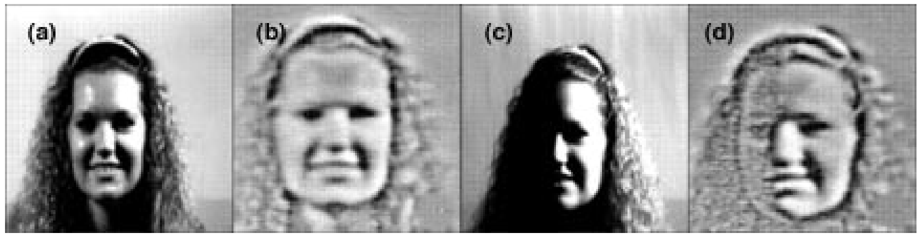
KEN Face Recognition Technology

University of California
Lawrence Livermore National Laboratory



The silicon retina extracts object information from the unilluminated face area that cannot be gathered by the CCD camera. Compare the images taken in two-sided illumination by the (a) CCD camera and (b) silicon retina with the pair taken in one-sided illumination by the (c) CCD camera and (d) silicon retina.

Objective

The recognition of human faces demonstrates object recognition under real-world conditions. It is based on wavelet-derived feature vectors and elastic matching. A silicon retina developed at California Institute of Technology can be used to reduce the system's dependency on lighting conditions.

Impact

In the era of the information superhighway, a wealth of application areas will need new technologies for object recognition, including active vision, robotics, security and surveillance, associative query of image databases, and human–computer interfaces.

The Institute for Scientific Computing Research (ISCR) at Lawrence Livermore National Laboratory recently developed the real-time recognition technology KEN. The name KEN is derived from the verb "to ken = to see, to recognize."

Handwriting and speech recognition are already part of today's multimedia PCs and personal digital assistants (PDAs). Object recognition, specifically face recognition, will be a natural extension of this technology. Associative query of image databases (e.g., search of medical image databases) and the interpretation of facial expressions for human–computer interfaces are likely applications. The existence of video-conferencing and the advent of interactive television widen the field of possible applications (e.g., model-based compression techniques for low-bit-rate video-conferencing).

In security and surveillance, applications include ID verification at automatic teller machines (ATMs), credit card verification, narrowing down a database of crime suspects by matching sketches, and identification of suspects for America's most wanted or suspected terrorists. The following scenarios are thinkable: An automatic recognition system preselects a few pictures of suspects from a crime scene out of a large database or alerts human security personnel if suspicious activity occurs around an ATM or a property. Valuable properties such as cars or boats could identify their owners. This recognition technology is not limited to faces but can also be applied to other object classes.

Recognition Process

KEN face recognition technology demonstrates object recognition technology by recognizing human faces in a real-world environment. The recognition technology has several modular components: The input sensor is either an off-the-shelf charge-coupled device (CCD) camera or an analog silicon retina chip (from Caltech). The silicon retina normalizes contrast locally and offers distinct advantages under strongly varying lighting conditions by extending the dynamic range beyond that of a normal CCD camera. A computer-controlled lens system with autozoom, autoiris, and autofocus and a pan-tilt unit can track possible face candidates. A Datacube real-time video-processing board extracts features and segments the input images down to a foveal area of attention. The control software for mechanical components and the classification software run on a workstation or PC.

KEN's software represents faces as labeled graphs. Wavelet-derived features encoding local object qualities are linked together into a graph model representing the topology of a face in its link structure. To "acquaint" the system with a person during the learning phase, the person's face is extracted as a frontal view and stored as a face graph model in the system's knowledge database. The knowledge base contains all memorized faces in the form of sparse graphs.

During recognition, face models are matched to the unknown face candidates appearing in the input. The graphs are positioned over a face candidate and normalized in size and orientation to the candidate by a template match. Subsequently, the graphs are locally distorted by an elastic matching process to fit the input image as closely as possible. The distortion effect is similar to the stretching of an elastic membrane (e.g., a rubber sheet). Cost values are assigned to the distorted face models according to the proximity of the final distorted graph to the input. A statistical evaluation of the resulting cost values of all models allows the system either to qualify a match or to reject poor matches and ambiguous results.

Rubber sheet effect: The labeled graph model (overlay) for the face in the left image is distorted to fit best over the face in the right image. The quality of the fit describes how closely the two face images resemble each other.

KEN is currently capable of classifying a face, almost in real time, from a database of 100 faces with a positive identification rate of up to 90% and a false positive recognition rate of less than 1%.

Environment

The hardware for KEN is commercially available except for the silicon retina prototype. The software is called AVision. It is written as a collection of object libraries in C and C++ and runs on major workstations (Sun, SGI, PC) and supercomputers, depending on the availability of C++ compilers. The upcoming development of offline learning algorithms for system optimization in a parallel implementation will require the performance of supercomputers.

KEN technology takes advantage of the World Wide Web technology by presenting an online face recognition demonstration that offers image format conversion and runs with an HTML 3.0–capable client (no http server required) under http://www-iscr.llnl.gov/KEN/KENOnline. A pilot implementation with JAVA interface is currently being tested.

Future Directions

We currently are researching an extension of this work to motion sequences of objects. This work involves constructing a choreographic motion database to implement a motion recognition system, since motion is used by humans as an additional visual cue. It appears to be an important component in a real-world application of recognition technology (e.g., crime scene detection). Other research is aimed at organizing large databases to improve the efficiency of the elastic matching process. Clustering processes, active data acquisition, and learning algorithms for the weighting of components are under development. KEN is well suited for the associative query of pictorial databases such as medical images.

Collaborations

ISCR is grateful to K. Boahen, T. Delbruck, and C. Mead, Caltech, for providing prototypes of a silicon retina. J. Buhmann of Bonn, Germany, contributed active vision concepts, learning algorithms, clustering schemes, and programming effort to the project. Additional collaborations include T. Sejnowski at the Salk Institute, for learning algorithms, and the University of California, San Diego and Los Angeles, for motion analysis.

We are currently seeking opportunities to collaborate with industry on security applications such as the verification of credit cards with pictures, the design of face recognition applications for the desktop, multimedia aspects of our recognition technology, and data collection for further development.

For more information, contact Martin Lades, 510-422-0893, hml@llnl.gov.

Work performed under the auspices of the U.S. DOE by LLNL under contract No. W-7405-ENG-48